

A hybrid entropic proximal decomposition method with self-adaptive strategy for solving variational inequality problems[☆]

Deren Han

Institute of Mathematics, School of Mathematics and Computer Science, Nanjing Normal University, Nanjing, 210097, China

Received 17 November 2006; received in revised form 1 February 2007; accepted 16 March 2007

Abstract

In this paper, we propose a hybrid nonlinear decomposition–projection method for solving a class of monotone variational inequality problems. The algorithm utilizes the problems' structure conducive to decomposition and a projection step to get the next iterate. To make the method more practical, we allow solving of the subproblems approximately and adopt the constructive accuracy criterion developed recently by Solodov and Svaiter for classical proximal point algorithm and by the author for generalized proximal point algorithm. The Fejér monotonicity to the solution set of the problem is obtained by only assuming the underlying mapping is monotone and the solution set is nonempty. The parameter is allowed to vary in a larger interval than that of Auslender and Teboulle, and we also propose some improved self-adaptive strategies to choose the sequence of parameters, which makes the algorithm more flexible.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Variational inequality problems; Decomposition methods; Entropic/interior proximal methods; Relative error criterion; Self-adaptive strategy

1. Introduction

Let Ω be a closed convex subset of R^n , F be a continuous mapping from R^n into itself, and $\langle \cdot, \cdot \rangle$ stands for the usual product in R^n . The variational inequality problem, which we abbreviate as $VIP(F, \Omega)$, is the problem of finding a vector $u^* \in \Omega$, such that

$$\langle u - u^*, F(u^*) \rangle \geq 0, \quad \forall u \in \Omega. \quad (1)$$

In this paper, we pay our attention to the $VIP(F, \Omega)$ with the following form

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, \quad F(u) = \begin{pmatrix} f(x) \\ g(y) \end{pmatrix}, \quad (2)$$

$$\Omega = \{(x, y) \mid Ax + By = b, x \in R_+^n, y \in R^m\}, \quad (3)$$

[☆] This research was supported by NSFC grant 10501024.

E-mail address: handeren@njnu.edu.cn.

where R_+^n is the nonnegative orthant of R^n , $A \in R^{l \times n}$ and $B \in R^{l \times m}$ are given matrices, $b \in R^l$ is a given vector, $f : R_+^n \rightarrow R^n$ and $g : R^m \rightarrow R^m$ are given monotone operators, respectively. Although problem (2) and (3) is a special case of the general VI problem, it has numerous important applications, especially in economics and transportation equilibrium problems [1–3]. For solving structured problem (2) and (3), a number of decomposition methods have been suggested in the literature, such as [1,3–7].

It is well known that $\text{VIP}(F, \Omega)$ is closely related to the problem of finding a zero of a maximal monotone operator T : find $u^* \in R^n$, such that

$$0 \in T(u^*). \quad (4)$$

Recall that an operator T is said to be monotone if for all $u_1, u_2 \in R^n$,

$$\langle u_1 - u_2, v_1 - v_2 \rangle \geq 0, \quad \forall v_1 \in T(u_1), v_2 \in T(u_2),$$

and T is said to be maximal monotone if the graph

$$G(T) = \{(u, v) \in R^n \times R^n \mid v \in T(u)\}$$

is not properly contained in the graph of any other monotone operators. If the relative interior of Ω intersects the interior of the domain of F , then the operator $T = F + N_\Omega$ is maximal monotone, and (1) can be reformulated as (4), where

$$N_\Omega(u) = \begin{cases} \{v : \langle u' - u, v \rangle \leq 0, \forall u' \in \Omega\}, & \text{if } u \in \Omega \\ \emptyset & \text{otherwise} \end{cases}$$

is the normal cone operator. Conversely, if $\Omega = R^n$, then (1) reduces to (4) with $T = F$.

A classic method for solving (4) is the proximal point algorithm [8,9]. By solving a class of subproblems

$$0 \in T_k(u), \quad T_k(u) = c_k T(u) + (u - u^k),$$

a sequence $\{u^k\}$ is generated, where u^k is the current approximation to a solution, and c_k is a positive regularization parameter. When applied to (1), the method needs to solve a sequence of variational inequality problems of the form

$$\text{VIP}(F_k, \Omega), \quad F_k(u) = c_k F(u) + (u - u^k).$$

Especially when applied to the problem (2) and (3), a class of decomposition methods such as the Douglas–Rachford splitting method and its variants were proposed [1,3,5,6,10–12]. Recently, generalized proximal point algorithms using nonlinear functions such as Bregman functions, φ -divergence and logarithmic–quadratic have been studied extensively [13–20]. A nonlinear proximal point algorithm consists of solving subproblems of the form

$$0 \in c_k T(u) + \Phi'(u^k, u), \quad (5)$$

where Φ' is a strictly monotone operator. When applied to variational inequality problems, the subproblems are essentially systems of equations, which are structurally considerable easier to solve than the variational inequality problems.

In [15], Auslender and Teboulle proposed an entropic decomposition method for solving convex programs and variational inequality problems with structure (2) and (3):

Let $\{c_k\}$ be a sequence of positive scalars and $\theta > 0$. Start with an arbitrary point $(x^0, y^0, \lambda^0) \in R_{++}^n \times R^m \times R^l$ and generate the sequence $\{x^k, y^k, \lambda^k\} \in R_{++}^n \times R^m \times R^l$ by the following steps:

Step 1. Compute

$$p^{k+1} = \lambda^k - (2\theta)^{-1} c_k (Ax^k + By^k - b). \quad (6)$$

Step 2. Find $(x^{k+1}, y^{k+1}) \in R_{++}^n \times R^m$ by solving

$$f(x^{k+1}) - A^\top p^{k+1} + c_k^{-1} \Phi'(x^k, x^{k+1}) = 0. \quad (7)$$

$$g(y^{k+1}) - B^\top p^{k+1} + 2\theta c_k^{-1} (y^{k+1} - y^k) = 0. \quad (8)$$

This method is an extension of the predictor–corrector proximal multiplier method developed by Chen and Teboulle [6]. They showed that under the assumption that

$$c_k \geq \eta > 0, \quad c_k \|A\| \leq \theta(\bar{\theta} - \epsilon)^{1/2}, \quad c_k \|B\| \leq \theta(1 - \epsilon)^{1/2}, \quad \forall k \geq 0, \quad (9)$$

the generated sequence converges to a solution of (2) and (3), where $\bar{\theta} = (v - \theta)/\theta$, v is a parameter used to define Φ' and $\epsilon \in (0, \theta)$. From (9), if c_k is chosen too large the method might not converge, while a choice for it that is too small might result in slowing the convergence. The choice of an approximate parameter c_k is thus crucial in any practical implementation of the algorithm.

Note that the main task of the method is to solve two systems of equations, which is structurally much easier to solve than variational inequalities, the subproblems of the classical decomposition algorithms [1,3,5–7,10,11,21,22]. However sometimes it could be very expensive or impossible to find the accurate solutions of (7) and (8). On the other hand, little justification has been provided to seek an accurate solution of (7) or (8) per iteration. In this paper, we propose a hybrid nonlinear proximal decomposition–projection method for solving these structured variational inequality problem (2) and (3). At each iteration, we first decompose the problem to two small problems with respect to x and y , respectively. Then, we take a projection step to generated the next iteration. This is what our method is named after. We show the convergence of the algorithm under the condition that the underlying mappings f and g are monotone and the parameter satisfies

$$0 < \underline{c} \leq c_k \leq \bar{c} < +\infty. \quad (10)$$

Note that this condition is much less stringent than (9), which allows us to choose a large ‘suitable’ parameter c_k . The self-adaptive rules we used here are the improved versions of the ones proposed in [21,23–26], where the summability of the adjustment parameter is relaxed to *any* positive constants, which means that we can adjust the parameter at *any time* when necessary. This may result in fast convergence and make the method more flexible, as shown in [21,23–26]. To make the method more practical, we allow solving of the subproblems approximately. The accuracy criterion we adopt here is just the one proposed in [27] for classical proximal point algorithms and in [28] for generalized proximal point algorithms, which is more constructive than the classical one assuming the summability of the parameters [6,9,29,30], and allows us to solve the subproblems via a few Newton-type steps.

Some other error criteria for the inexact version of the generalized proximal point algorithm (5)

$$e^{k+1} \in c_k T(u^{k+1}) + \Phi'(u^k, u^{k+1}), \quad (11)$$

are [18,20,31,32]. Among them, Eckstein’s criterion [18]

$$\begin{aligned} \sum_{k=0}^{\infty} \|e^k\| < \infty \quad \text{and} \\ \sum_{k=0}^{\infty} \langle e^k, u^k \rangle \quad \text{exists and is finite} \end{aligned} \quad (12)$$

is the simplest and easiest to use in practice. Still, the approximate criterion (12) is more restrictive than that for classical proximal point algorithm, which only requires the first inequality in (12). Recently, Solodov and Svaiter [19] proposed a new generalized proximal point algorithm. At the k th step, for given u^k , they get a proximal solution $u^{k+1} \in \Omega$ by solving

$$0 \in c_k T(v^k) + \Phi'(u^k, v^k) \quad (13)$$

and

$$u^{k+1} = \Phi'(u^k, \cdot)^{-1}(-c_k T(v^k)) \quad (14)$$

satisfying

$$d(v^k, u^{k+1}) \leq \sigma^2 d(v^k, u^k), \quad (15)$$

where $d(\cdot, \cdot)$ is “ D -function” which will be defined in the sequel and $\sigma \in (0, 1)$ is a constant. Note that the error tolerance (15) is more constructive than (12), because $\sigma \in (0, 1)$ is a constant. However, to verify if v^k is an acceptable proximal solution, we have first to solve the problem (14)

$$\Phi'(u^k, u) = -c_k T(v^k),$$

to get a trial point u^{k+1} , which may be computationally expensive in many cases when Φ' is difficult to invert. Their method, therefore, has an advantage only for the case that $\Phi'(u^k, \cdot)^{-1}$ is easy to get. The most recent one of Burachik and Svaiter [33] shares the same limitation as the one of Solodov and Svaiter [19].

The rest of the paper is organized as follows. In Section 2, we state the definition and summarize some properties of logarithmic–quadratic functions. In Section 3, the new algorithm is described. In Section 4, we prove the global convergence of the proposed method. Two improved adjustment strategies are discussed in Section 5, and we conclude the paper by giving some remarks in Section 6.

2. Logarithmic–quadratic function

In this section, we state the definition and summarize some properties of logarithmic–quadratic functions that will be used in the following discussion.

Let $\nu > \mu > 0$ be given fixed parameters and define the logarithmic–quadratic function as

$$\varphi(t) = \begin{cases} \frac{\nu}{2}(t-1)^2 + \mu(t - \log t - 1) & \text{if } t > 0 \\ +\infty & \text{otherwise.} \end{cases} \quad (16)$$

From the definition of φ , we have that φ is a differentiable strongly convex function on R_{++}^n with modulus $\nu > 0$ and

$$\lim_{t \rightarrow 0^+} \varphi'(t) = -\infty.$$

Associated with φ we define for any $v \in R_{++}^n$

$$d(u, v) = \begin{cases} \sum_{i=1}^n \left(\frac{\nu}{2}(u_i - v_i)^2 + \mu \left(v_i^2 \log \frac{v_i}{u_i} + u_i v_i - v_i^2 \right) \right) & \text{if } u \in R_{++}^n \\ +\infty & \text{otherwise.} \end{cases}$$

For any $u, v \in R_{++}^n$, the above defined functional d can be rewritten as

$$d(u, v) = \sum_{i=1}^n v_i^2 \varphi(u_i v_i^{-1}).$$

It follows from the definitions of φ and d that for any $u, v \in R_{++}^n$

$$d(u, v) \geq \frac{\nu}{2} \|u - v\|^2 \quad (17)$$

and

$$d(u, v) = 0 \iff u = v.$$

For convenience, we set $\theta := (\nu + \mu)/2$ and $\bar{\theta} := (\nu - \theta)/\theta$. We use the following notation in the rest of the paper

$$\Phi'(a, b) := (a_1 \varphi'(b_1/a_1), \dots, a_n \varphi'(b_n/a_n))^T, \quad \forall a, b \in R_{++}^n. \quad (18)$$

The following result was recently derived in [13].

Lemma 2.1. *Let φ be given in (16) and Φ' be defined in (18). For any $a, b \in R_{++}^n$ and $c \in R_+^n$, we have*

$$\langle c - b, \Phi'(a, b) \rangle \leq \theta(\|c - a\|^2 - \|c - b\|^2 - \bar{\theta}\|b - a\|^2). \quad (19)$$

From this lemma, it is easy to derive the following result (just by replacing c with a in the above lemma).

Lemma 2.2. Let φ be given in (16) and Φ' be defined in (18). For any $a, b \in R_{++}^n$ we have

$$\langle b - a, \Phi'(a, b) \rangle \geq v \|a - b\|^2. \quad (20)$$

Let $P_X[a]$ denote the projection of a vector $a \in R^n$ on $X = R_+^n$, i.e., $P_X[a] = \max\{a, 0\}$. Then for any $a, b \in R^n$

$$\|P_X[a] - P_X[b]\| \leq \|a - b\|. \quad (21)$$

3. The hybrid algorithm

In this section, we propose the hybrid algorithm for solving variational inequality problems with the structure (2) and (3). This algorithm can be viewed as an extension of the method [28] to the separable problem (2) and (3).

We make the following standard assumption for the variational inequality problem $VI(F, \Omega)$:

Assumption A. (a) Problem $VI(F, \Omega)$ has a solution;

(b) $\text{dom } T \cap (R_{++}^n \times R^m) \neq \emptyset$,

where $\text{dom } T$ denotes the domain of the operator T , which is defined as

$$\text{dom } T := \{x \mid T(x) \neq \emptyset\}.$$

Then, it is well known that (x^*, y^*) solves $VI(F, \Omega)$ if and only if there exists a $\lambda^* \in R^l$, multiplier for the constraint $Ax + By = b$ such that (x^*, y^*, λ^*) solves the following variational inequality problem: Find $w^* \in W$, such that

$$\langle Q(w^*), w - w^* \rangle \geq 0, \quad \forall w \in W, \quad (22)$$

where

$$w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, \quad Q(w) = \begin{pmatrix} f(x) - A^\top \lambda \\ g(y) - B^\top \lambda \\ Ax + By - b \end{pmatrix}, \quad W = R_+^n \times R^m \times R^l. \quad (23)$$

For convenience, we denote VI problem (22) and (23) by $MVI(Q, W)$. In the following of this paper, instead of solving the primal problem (2) and (3), we will give our attention to solving the equivalent problem $MVI(Q, W)$.

Algorithm 1 (Hybrid Decomposition–Projection Algorithm).

Step 0. Choose some $0 < \underline{c} \leq \bar{c} < +\infty$, $t \in (0, 1)$, the error tolerance parameter $\sigma \in (0, \min\{1, v\})$, and $(x^0, y^0, \lambda^0) \in R_{++}^n \times R^m \times R^l$. Set $k := 0$.

Step 1. Choose the regularization parameter $c_k \in [1 - \sigma, \bar{c}]$, find an inexact solution $\bar{x}^k \in R_{++}^n$ of the proximal subproblem

$$c_k(f(\cdot) - A^\top \lambda^k) + \Phi'(x^k, \cdot) = 0,$$

satisfying

$$\|r_x^k\| \leq \sigma \|x^k - \bar{x}^k\|, \quad (24)$$

where

$$c_k(f(\cdot) - A^\top \lambda^k) + \Phi'(x^k, \cdot) = r_x^k. \quad (25)$$

Step 2. Find an inexact solution $\bar{y}^k \in R^m$ of the proximal subproblem

$$c_k(g(\cdot) - B^\top \lambda^k) + (\cdot - y^k) = 0$$

satisfying

$$\|r_y^k\| \leq \sigma \|y^k - \bar{y}^k\|, \quad (26)$$

where

$$c_k(g(\cdot) - B^\top \lambda^k) + (\cdot - y^k) = r_y^k. \quad (27)$$

Step 3. Set

$$p^k = \lambda^k - (Ax^k + By^k - b),$$

and compute x^{k+1}, y^{k+1} via

$$x^{k+1} = (1-t)P_X[x^k - \alpha_k(f(\bar{x}^k) - A^\top p^k)] + tx^k, \quad (28)$$

$$y^{k+1} = y^k - (1-t)\alpha_k(g(\bar{y}^k) - B^\top p^k), \quad (29)$$

where

$$\alpha_k = \zeta_k / \xi_k \quad (30)$$

and

$$\zeta_k = \langle f(\bar{x}^k) - A^\top \lambda^k, x^k - \bar{x}^k \rangle + \langle g(\bar{y}^k) - B^\top \lambda^k, y^k - \bar{y}^k \rangle + \|Ax^k + By^k - b\|^2, \quad (31)$$

$$\xi_k = \|f(\bar{x}^k) - A^\top p^k\|^2 + \|g(\bar{y}^k) - B^\top p^k\|^2 + (1-t)\|A\bar{x}^k + B\bar{y}^k - b\|^2. \quad (32)$$

Step 4. Compute λ^{k+1}

$$\lambda^{k+1} = \lambda^k - (1-t)\alpha_k(A\bar{x}^k + B\bar{y}^k - b). \quad (33)$$

Step 5. Choose a parameter c_{k+1} for the next step. Set $k := k + 1$ and go to Step 1.

Remark 3.1. In the above algorithm, we have not specified the rules on how to choose c_k and the concrete description will be given in Section 5. Here, we just require $c_k \in [1 - \sigma, \bar{c}] \subset (0, \infty)$, and any strategy satisfying this condition can be used in Step 5.

Lemma 3.1. For any $c_k > 0$, $(x^k, y^k, \lambda^k) \in R_{++}^n \times R^m \times R^l$ and $\forall k \geq 0$, there exists a unique point $(x, y) \in R_{++}^n \times R^m$ satisfying

$$c_k(f(\cdot) - A^\top \lambda^k) + \Phi'(x^k, \cdot) = 0,$$

and

$$c_k(g(\cdot) - B^\top \lambda^k) + (\cdot - y^k) = 0.$$

Proof. This is an immediate consequence of Proposition 2 in [14]. \square

The above lemma guarantees that the generalized proximal subproblems (25) and (27) always have exact solutions in R_{++}^n ($r_x^k = 0$) and R^m ($r_y^k = 0$). These two problems certainly always have inexact solutions \bar{x}^k and \bar{y}^k satisfying $\bar{x}^k \in R_{++}^n$, $\bar{y}^k \in R^m$, respectively. The proximal subproblems are thus well defined for all $k \geq 0$. Hence, we can conclude that the whole algorithm is well-defined. Moreover, since $P_X(\cdot) \in R_+^n$, the generated sequence $\{(x^k, y^k)\}$ is contained in $R_{++}^n \times R^m$.

In Step 1, to find the solution \bar{x}^k , one can solve the equation

$$c_k(f(\cdot) - A^\top \lambda^k) + \Phi'(x^k, \cdot) = 0 \quad (34)$$

by Newton's method [34] (with starting point $x_0 := x^k$), and stop with the first Newton iterate satisfying (24). Since f is continuous and monotone, the mapping $\Phi' + c_k f$ is strongly monotone. This system of nonlinear equations is well conditioned. Moreover, since for k large enough, x^k is close to \bar{x}^k (see the following proof), Newton's method can find a solution of this equation within finitely many iterations

4. Global convergence

In this section, we will show that the proposed algorithm converges to a solution of the variational inequality problem (2) and (3) globally. In the remainder of this section, we always suppose that (22) and (23) has at least one solution $w^* \in W$.

We now begin to analyze the convergence of the proposed algorithm with the following result, which is a classical estimate in proximal point algorithms, see e.g. [35].

Lemma 4.1. Let $f : \rightarrow (-\infty, +\infty]$ be a closed proper convex function, and $\tilde{s} \in R^n$ be given by

$$\tilde{s} = \arg \min_{s \in R^n} \left\{ f(s) + \frac{1}{2c} \|s - \tilde{s}\|^2 \right\},$$

where $c > 0$. Then, for any $s \in R^n$, the following inequality holds:

$$2c\{f(\tilde{s}) - f(s)\} \leq \|\tilde{s} - s\|^2 - \|\tilde{s} - s\|^2 - \|\tilde{s} - \tilde{s}\|^2. \quad (35)$$

This lemma states a useful relation between two consecutive iterates generated by the hybrid proximal decomposition algorithms.

Lemma 4.2. Let $\{x^k\}$, $\{y^k\}$, $\{\bar{x}^k\}$, $\{\bar{y}^k\}$ and $\{\lambda^k\}$ be generated by Algorithm 1. Then, for all $(x, y, \lambda) \in R_+^n \times R^m \times R^l$, the following inequalities hold:

$$\|x^{k+1} - x\|^2 \leq \|x^k - x\|^2 + (1-t)\alpha_k^2 \|f(\bar{x}^k) - A^\top p^k\|^2 - 2(1-t)\alpha_k \langle f(\bar{x}^k) - A^\top p^k, x^k - x \rangle, \quad (36)$$

$$\|y^{k+1} - y\|^2 \leq \|y^k - y\|^2 + (1-t)\alpha_k^2 \|g(\bar{y}^k) - B^\top p^k\|^2 - 2(1-t)\alpha_k \langle g(\bar{y}^k) - B^\top p^k, y^k - y \rangle \quad (37)$$

$$\|\lambda^{k+1} - \lambda\|^2 \leq \|\lambda^k - \lambda\|^2 - 2\alpha_k(1-t) \langle \lambda^{k+1} - \lambda, A\bar{x}^k + B\bar{y}^k - b \rangle - \|\lambda^{k+1} - \lambda^k\|^2. \quad (38)$$

Proof. It follows from (28) that

$$\begin{aligned} \|x^{k+1} - x\|^2 &= \|(1-t)(P_X[x^k - \alpha_k(f(\bar{x}^k) - A^\top p^k)] - x) + t(x^k - x)\|^2 \\ &= (1-t)^2 \|P_X[x^k - \alpha_k(f(\bar{x}^k) - A^\top p^k)] - x\|^2 + t^2 \|x^k - x\|^2 \\ &\quad + 2t(1-t) \langle P_X[x^k - \alpha_k(f(\bar{x}^k) - A^\top p^k)] - x, x^k - x \rangle. \end{aligned}$$

Since $2\langle a, b \rangle \leq \|a\|^2 + \|b\|^2$, we have

$$\begin{aligned} &2t(1-t) \langle P_X[x^k - \alpha_k(f(\bar{x}^k) - A^\top p^k)] - x, x^k - x \rangle \\ &\leq t(1-t) \|P_X[x^k - \alpha_k(f(\bar{x}^k) - A^\top p^k)] - x\|^2 + t(1-t) \|x^k - x\|^2. \end{aligned}$$

Thus,

$$\begin{aligned} \|x^{k+1} - x\|^2 &\leq (1-t)^2 \|P_X[x^k - \alpha_k(f(\bar{x}^k) - A^\top p^k)] - x\|^2 + t^2 \|x^k - x\|^2 \\ &\quad + t(1-t) (\|P_X[x^k - \alpha_k(f(\bar{x}^k) - A^\top p^k)] - x\|^2 + \|x^k - x\|^2) \\ &= (1-t) \|P_X[x^k - \alpha_k(f(\bar{x}^k) - A^\top p^k)] - x\|^2 + t \|x^k - x\|^2. \end{aligned}$$

Since the projection operator P_X is nonexpansive and $x \in R_+^n$, we have

$$\begin{aligned} \|x^{k+1} - x\|^2 &\leq (1-t) \|P_X[x^k - \alpha_k(f(\bar{x}^k) - A^\top p^k)] - x\|^2 + t \|x^k - x\|^2 \\ &\leq \|x^k - x\|^2 - 2(1-t)\alpha_k \langle f(\bar{x}^k) - A^\top p^k, x^k - x \rangle + (1-t)\alpha_k^2 \|f(\bar{x}^k) - A^\top p^k\|^2, \end{aligned} \quad (39)$$

which is just (36). This completes the proof of the inequality (36), and the inequality (37) can be proved in a similar way.

Observe that (33) implies

$$\lambda^{k+1} = \arg \min_{\lambda \in R^l} \left\{ \langle \lambda, A\bar{x}^k + B\bar{y}^k - b \rangle + \frac{1}{2(1-t)\alpha_k} \|\lambda - \lambda^k\|^2 \right\}.$$

From Lemma 4.1, we have that

$$2(1-t)\alpha_k \langle \lambda^{k+1} - \lambda, A\bar{x}^k + B\bar{y}^k - b \rangle \leq \|\lambda^k - \lambda\|^2 - \|\lambda^{k+1} - \lambda\|^2 - \|\lambda^{k+1} - \lambda^k\|^2.$$

By rearranging terms, we obtain (38) immediately. \square

Lemma 4.3. Suppose that f , g are continuous and monotone, and $w^* = (x^*, y^*, \lambda^*) \in W$ is a solution of MVI(Q , W). Then

1. The generated sequence $\{w^k\}$ is bounded;
2. The sequence $\{(\bar{x}^k, \bar{y}^k)\}$ is bounded;
3. $\lim_{k \rightarrow \infty} (x^k - \bar{x}^k) = \lim_{k \rightarrow \infty} (y^k - \bar{y}^k) = \lim_{k \rightarrow \infty} \|Ax^k + By^k - b\| = 0$.

Proof. Since f and g are monotone mappings, we have that for any $w_1 = (x_1, y_1, \lambda_1)$ and $w_2 = (x_2, y_2, \lambda_2)$

$$\begin{aligned} (w_1 - w_2)^\top (Q(w_1) - Q(w_2)) &= \begin{pmatrix} x_1 - x_2 \\ y_1 - y_2 \\ \lambda_1 - \lambda_2 \end{pmatrix}^\top \begin{pmatrix} f(x_1) - f(x_2) - A^\top(\lambda_1 - \lambda_2) \\ g(y_1) - g(y_2) - B^\top(\lambda_1 - \lambda_2) \\ A(x_1 - x_2) + B(y_1 - y_2) \end{pmatrix} \\ &= (x_1 - x_2)^\top (f(x_1) - f(x_2)) + (y_1 - y_2)^\top (g(y_1) - g(y_2)) \\ &\geq 0. \end{aligned}$$

That is, Q defined by (23) is a monotone mapping from W to $R^n \times R^m \times R^l$. Since (x^*, y^*, λ^*) is a solution of MVI(Q, W) and $\bar{x}^k \in R_+^n, \bar{y}^k \in R^m$, we have

$$\begin{pmatrix} f(\bar{x}^k) - A^\top \lambda^k \\ g(\bar{y}^k) - B^\top \lambda^k \\ A\bar{x}^k + B\bar{y}^k - b \end{pmatrix}^\top \begin{pmatrix} \bar{x}^k - x^* \\ \bar{y}^k - y^* \\ \lambda^k - \lambda^* \end{pmatrix} \geq \begin{pmatrix} f(x^*) - A^\top \lambda^* \\ g(y^*) - B^\top \lambda^* \\ Ax^* + By^* - b \end{pmatrix}^\top \begin{pmatrix} \bar{x}^k - x^* \\ \bar{y}^k - y^* \\ \lambda^k - \lambda^* \end{pmatrix} \geq 0,$$

which means that

$$\langle f(\bar{x}^k) - A^\top \lambda^k, \bar{x}^k - x^* \rangle + \langle g(\bar{y}^k) - B^\top \lambda^k, \bar{y}^k - y^* \rangle \geq -\langle \lambda^k - \lambda^*, A\bar{x}^k + B\bar{y}^k - b \rangle. \quad (40)$$

Using the identity

$$Ax^* + By^* = b,$$

it follows from (31) and (40) that

$$\langle f(\bar{x}^k) - A^\top p^k, x^k - x^* \rangle + \langle g(\bar{y}^k) - B^\top p^k, y^k - y^* \rangle \geq \zeta_k - \langle \lambda^k - \lambda^*, A\bar{x}^k + B\bar{y}^k - b \rangle. \quad (41)$$

Adding (36) and (37) and substituting (x^*, y^*) for (x, y) , we have

$$\begin{aligned} &\|x^{k+1} - x^*\|^2 + \|y^{k+1} - y^*\|^2 \\ &\leq \|x^k - x^*\|^2 + \|y^k - y^*\|^2 + (1-t)\alpha_k^2 \{\|f(\bar{x}^k) - A^\top p^k\|^2 + \|g(\bar{y}^k) - B^\top p^k\|^2\} \\ &\quad - 2(1-t)\alpha_k \zeta_k + 2(1-t)\alpha_k \langle \lambda^k - \lambda^*, A\bar{x}^k + B\bar{y}^k - b \rangle, \end{aligned} \quad (42)$$

where the inequality follows from (41). Substituting λ by λ^* in (38) and using (33), we have

$$\|\lambda^{k+1} - \lambda^*\|^2 \leq \|\lambda^k - \lambda^*\|^2 - 2\alpha_k(1-t)\langle \lambda^{k+1} - \lambda^*, A\bar{x}^k + B\bar{y}^k - b \rangle - \alpha_k^2(1-t)^2 \|A\bar{x}^k + B\bar{y}^k - b\|^2. \quad (43)$$

Adding (43) to (42), it follows that

$$\begin{aligned} &\|x^{k+1} - x^*\|^2 + \|y^{k+1} - y^*\|^2 + \|\lambda^{k+1} - \lambda^*\|^2 \\ &\leq \|x^k - x^*\|^2 + \|y^k - y^*\|^2 + \|\lambda^k - \lambda^*\|^2 - 2(1-t)\alpha_k \zeta_k \\ &\quad + (1-t)\alpha_k^2 \{\|f(\bar{x}^k) - A^\top p^k\|^2 + \|g(\bar{y}^k) - B^\top p^k\|^2 + (1-t)\|A\bar{x}^k + B\bar{y}^k - b\|^2\} \\ &= \|x^k - x^*\|^2 + \|y^k - y^*\|^2 + \|\lambda^k - \lambda^*\|^2 - (1-t)\alpha_k \zeta_k, \end{aligned} \quad (44)$$

where the last equality follows from the definition of α_k . Note that

$$f(\bar{x}^k) - A^\top \lambda^k = \frac{1}{c_k} [r_x^k - \Phi'(x^k, \bar{x}^k)]. \quad (45)$$

Thus, it follows from Lemma 2.2 that

$$\begin{aligned} \langle f(\bar{x}^k) - A^\top \lambda^k, x^k - \bar{x}^k \rangle &= \frac{1}{c_k} \langle r_x^k - \Phi'(x^k, \bar{x}^k), x^k - \bar{x}^k \rangle \\ &\geq \frac{1}{c_k} (\langle r_x^k, x^k - \bar{x}^k \rangle + \theta(1+\bar{\theta})\|x^k - \bar{x}^k\|^2) \end{aligned}$$

$$\begin{aligned}
&\geq \frac{1}{c_k}(\theta(1 + \bar{\theta})\|x^k - \bar{x}^k\|^2 - \|r_x^k\|\|x^k - \bar{x}^k\|) \\
&\geq \frac{1}{c_k}(\theta(1 + \bar{\theta}) - \sigma)\|x^k - \bar{x}^k\|^2,
\end{aligned} \tag{46}$$

where the second inequality follows from the Cauchy–Schwarz inequality, and the last one follows from (24). In a similar way, from (26), we have

$$\langle g(\bar{y}^k) - B^\top \lambda^k, y^k - \bar{y}^k \rangle \geq \frac{1}{c_k}(1 - \sigma)\|y^k - \bar{y}^k\|^2. \tag{47}$$

Let

$$\gamma := \min\{1, \theta(1 + \bar{\theta})\} \geq \sigma.$$

Then, substituting (46) and (47) in (44), we have

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 + \|y^{k+1} - y^*\|^2 + \|\lambda^{k+1} - \lambda^*\|^2 &\leq \|x^k - x^*\|^2 + \|y^k - y^*\|^2 + \|\lambda^k - \lambda^*\|^2 \\
&\quad - (1 - t)\alpha_k \left[\frac{(\gamma - \sigma)}{c_k} (\|x^k - \bar{x}^k\|^2 \right. \\
&\quad \left. + \|y^k - \bar{y}^k\|^2) + \|Ax^k + By^k - b\|^2 \right].
\end{aligned} \tag{48}$$

Since $t \in [0, 1]$, $\sigma \in [0, \gamma]$ and $\alpha_k \geq 0$, we have

$$\|x^{k+1} - x^*\|^2 + \|y^{k+1} - y^*\|^2 + \|\lambda^{k+1} - \lambda^*\|^2 \leq \|x^k - x^*\|^2 + \|y^k - y^*\|^2 + \|\lambda^k - \lambda^*\|^2,$$

which means that the sequence $\{w^k\} = \{(x^k, y^k, \lambda^k)\}$ is bounded.

From the monotonicity of f and (46), we have

$$\langle f(x^k) - A^\top \lambda^k, x^k - \bar{x}^k \rangle \geq \langle f(\bar{x}^k) - A^\top \lambda^k, x^k - \bar{x}^k \rangle \geq \frac{1}{c_k}(\gamma - \sigma)\|x^k - \bar{x}^k\|^2.$$

Using again the Cauchy–Schwarz inequality, we have

$$\|f(x^k) - A^\top \lambda^k\| \geq \frac{1}{c_k}(\gamma - \sigma)\|x^k - \bar{x}^k\|.$$

Since $\{x^k\}$, $\{\lambda^k\}$ are bounded and f is continuous, $\{\bar{x}^k\}$ is bounded. The boundedness of $\{\bar{y}^k\}$ can be proved in a similar way.

Since f and g are continuous and the sequences $\{w^k\} = \{(x^k, y^k, \lambda^k)\}$, $\{(\bar{x}^k, \bar{y}^k)\}$ are bounded, there exist constants $M_1 > 0$, $M_2 > 0$ and $M_3 > 0$, such that

$$\begin{aligned}
\|f(\bar{x}^k) - A^\top p^k\| &\leq M_1, \\
\|g(\bar{y}^k) - B^\top p^k\| &\leq M_2,
\end{aligned}$$

and

$$\|A\bar{x}^k + B\bar{y}^k - b\| \leq M_3.$$

From the definition of α_k and (46) and (47), we have

$$\begin{aligned}
\alpha_k &\geq \frac{(\gamma - \sigma)(\|x^k - \bar{x}^k\|^2 + \|y^k - \bar{y}^k\|^2) + c_k\|Ax^k + By^k - b\|^2}{c_k(M_1^2 + M_2^2 + (1 - t)M_3^2)} \\
&\geq \frac{(\gamma - \sigma)(\|x^k - \bar{x}^k\|^2 + \|y^k - \bar{y}^k\|^2) + \underline{c}\|Ax^k + By^k - b\|^2}{\bar{c}(M_1^2 + M_2^2 + (1 - t)M_3^2)}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 + \|y^{k+1} - y^*\|^2 + \|\lambda^{k+1} - \lambda^*\|^2 &\leq \|x^k - x^*\|^2 + \|y^k - y^*\|^2 + \|\lambda^k - \lambda^*\|^2 \\
&\quad - \frac{(1 - t)}{\bar{c}^2(M_1^2 + M_2^2 + (1 - t)M_3^2)} \left[(\gamma - \sigma)^2(\|x^k - \bar{x}^k\|^2 + \|y^k - \bar{y}^k\|^2) + \underline{c}\|Ax^k + By^k - b\|^2 \right]^2.
\end{aligned}$$

It thus follows immediately that

$$\lim_{k \rightarrow \infty} \|x^k - \bar{x}^k\| = \lim_{k \rightarrow \infty} \|y^k - \bar{y}^k\| = \lim_{k \rightarrow \infty} \|Ax^k + By^k - b\| = 0.$$

This completes the proof. \square

Lemma 4.4. Suppose that f and g are continuous and monotone mappings and the sequence $\{w^k\}$ has a cluster point w^∞ , then w^∞ solves the variational inequality problem $MVI(Q, W)$.

Proof. It follows from (24) and (26) and Lemma 4.3 that $r_x^k \rightarrow 0$ and $r_y^k \rightarrow 0$. Let $\{w^{k_j}\}$ be the subsequence that converges to w^∞ . Then, from Lemma 4.3 we have that $\bar{x}^{k_j} \rightarrow x^\infty$ and $\bar{y}^{k_j} \rightarrow y^\infty$.

For any $x \in R_{++}^n$, it follows from (45) and Lemma 2.1 that

$$\begin{aligned} \langle f(\bar{x}^{k_j}) - A^\top \lambda^{k_j}, x - \bar{x}^{k_j} \rangle &= \frac{1}{c_{k_j}} \langle r_x^{k_j} - \Phi'(x^{k_j}, \bar{x}^{k_j}), x - \bar{x}^{k_j} \rangle \\ &\geq \frac{1}{c_{k_j}} \langle r_x^{k_j}, x - \bar{x}^{k_j} \rangle - \frac{\theta}{c_{k_j}} (\|x - x^{k_j}\|^2 - \|x - \bar{x}^{k_j}\|^2 - \bar{\theta} \|x^{k_j} - \bar{x}^{k_j}\|^2). \end{aligned}$$

Taking limits on both sides of the above inequality along the subsequence, we have

$$\langle f(x^\infty) - A^\top \lambda^\infty, x - x^\infty \rangle \geq 0.$$

In a similar way, we can prove that

$$\langle g(y^\infty) - B^\top \lambda^\infty, y - y^\infty \rangle \geq 0.$$

On the other hand, it follows from Lemma 4.3 that

$$Ax^\infty + By^\infty - b = 0.$$

Thus,

$$\begin{pmatrix} f(x^\infty) - A^\top \lambda^\infty \\ g(y^\infty) - B^\top \lambda^\infty \\ Ax^\infty + By^\infty - b \end{pmatrix}^\top \begin{pmatrix} x - x^\infty \\ y - y^\infty \\ \lambda - \lambda^\infty \end{pmatrix} \geq 0, \quad \forall (x, y, \lambda) \in W,$$

which means that w^∞ is a solution of $MVI(Q, W)$. \square

Theorem 4.5. Suppose that f and g are continuous and monotone and the solutions set of $MVI(Q, W)$ is nonempty. Then, the sequence $\{w^k\}$ generated by the hybrid method converges to a solution of $MVI(Q, W)$.

Proof. By Lemma 4.3, $\{w^k\}$, conforming to the algorithm, is bounded. It thus has at least one cluster point. Let \tilde{w} be a cluster point of $\{w^k\}$ and $\{w^{k_j}\}$ be the corresponding subsequence converging to \tilde{w} . Since $\{w^k\}$ contained in W , we have $\tilde{w} \in \bar{W}$. It follows from Lemma 4.4 that \tilde{w} solves $MVI(Q, W)$. Substitute (x^*, y^*, λ^*) in Lemma 4.3 and by $(\tilde{x}, \tilde{y}, \tilde{\lambda})$, we have

$$\|x^{k+1} - \tilde{x}\|^2 + \|y^{k+1} - \tilde{y}\|^2 + \|\lambda^{k+1} - \tilde{\lambda}\|^2 \leq \|x^k - \tilde{x}\|^2 + \|y^k - \tilde{y}\|^2 + \|\lambda^k - \tilde{\lambda}\|^2.$$

Thus, the whole sequence $\{w^k\}$ converges to \tilde{w} , a solution of $MVI(Q, W)$. \square

5. Choose the parameter

In a practical implementation of the algorithm, the choice of an appropriate parameter c_k is important, for on the one hand, if the parameter is set to be too small, the algorithm will converge slowly, and on the other hand, if c_k is too large, it may cause the divergence of the generated sequence. In practical applications, to choose a ‘suitable’ parameter c is difficult and many computational results [21,23–26] indicate that the method with a fixed parameter converges extremely slowly due to the parameter being too small. Although we just require the boundedness of the sequence $\{c_k\} \subseteq [c, \bar{c}] \subseteq (0, \infty)$, which is much less stringent than that of Auslender and Teboulle [15] (see also (9)), it seems reasonable to choose c_k self-adaptively per iteration.

Note that in Step 1 and Step 2, we need to solve approximately the system of equations of

$$c_k(f(\cdot) - A^\top \lambda^k) + \Phi'(x^k, \cdot) = 0$$

and

$$c_k(g(\cdot) - B^\top \lambda^k) + (\cdot - y^k) = 0,$$

respectively. For the sake of balance, it is reasonable to hope that

$$\left\| \begin{matrix} x^{k+1} - x^k \\ y^{k+1} - y^k \end{matrix} \right\| \approx c_k \left\| \begin{matrix} f(x^{k+1}) - f(x^k) \\ g(y^{k+1}) - g(y^k) \end{matrix} \right\|.$$

This consideration provides a way of choosing the parameter c_k , and in a similar way to He, Liao and Wang [26], we can use the following strategy:

Self-adaptive Strategy 1: From the viewpoint of subproblems

Let

$$\omega_k = \left\| \begin{matrix} x^{k+1} - x^k \\ y^{k+1} - y^k \end{matrix} \right\| / c_k \left\| \begin{matrix} f(x^{k+1}) - f(x^k) \\ g(y^{k+1}) - g(y^k) \end{matrix} \right\|, \quad (49)$$

Then, let

$$c_{k+1} = \begin{cases} \min\{(1 + \tau_k)c_k, \bar{c}\} & \text{if } \omega_k < \frac{1}{(1 + \mu)}, \\ \max\{\underline{c}, \frac{1}{(1 + \tau_k)}c_k\} & \text{if } \omega_k > 1 + \mu, \\ c_k & \text{otherwise,} \end{cases} \quad (50)$$

where $\mu > 0$ is a given constant.

Remark 5.1. In He, Liao and Wang [26], they need the condition that the nonnegative sequence $\{\tau_k\}$ satisfies $\sum_{k \rightarrow \infty} \tau_k < \infty$ to prove the global convergence of their method. Note that the above requirement is quite stringent, which means that we can just increase or decrease c_k by finite times. Here, the adjustment parameter τ_k can be a positive constant, due to the fact that we just need the boundedness of $\{c_k\}$. This is quite desirable, since it means that we can choose a suitable parameter by increasing or decreasing it whenever necessary.

Recall that instead of solving the original problem (1)–(3), we solve its equivalent form (22) and (23). It is well known (see for example, [36]) that solving (22) and (23) is equivalent to finding a zero point of $E(w, c)$:

$$E(w, c_k) := \begin{pmatrix} E_x(w, c_k) \\ E_y(w, c_k) \\ E_\lambda(w, c_k) \end{pmatrix} = \begin{pmatrix} x - P_X[x - c_k * (f(x) - A^\top \lambda)] \\ c_k * (g(y) - B^\top \lambda) \\ Ax + By - b \end{pmatrix}.$$

It follows from the nonexpansivity of the projection operator, that we have that

$$\|E(w, c_k)\| \leq \left\| \begin{matrix} c_k * (f(x) - A^\top \lambda) \\ c_k * (g(y) - B^\top \lambda) \\ Ax + By - b \end{matrix} \right\|.$$

Let

$$E_u(w, c) = \begin{pmatrix} E_x(w, c) \\ E_y(w, c) \end{pmatrix},$$

and note that for any fixed w , $E_u(w, \cdot)$ is an increasing function for the second parameter [37]. At the same time, $E_\lambda(w, c)$ is irrelative to c . Therefore, for the sake of balance, He, Liao, Han and Yang [21] suggested to adjust c_k with the following rule:

Self-adaptive Strategy 2: From the viewpoint of the original problem

Let

$$\omega_k = \|E_u(w^k, c_k)\| / \|E_\lambda(w^k, c_k)\|, \quad (51)$$

Then, adjust c_k in a similar way as Strategy 1:

$$c_{k+1} = \begin{cases} \min\{(1 + \tau_k)c_k, \bar{c}\} & \text{if } \omega_k < \frac{1}{(1 + \mu)}, \\ \max\{\underline{c}, \frac{1}{(1 + \tau_k)}c_k\} & \text{if } \omega_k > 1 + \mu, \\ c_k & \text{otherwise,} \end{cases} \quad (52)$$

where $\mu > 0$ is a given constant.

Remark 5.2. Note that in He, Liao, Han and Yang [21], they also need the condition that the nonnegative sequence $\{\tau_k\}$ satisfies $\sum_{k \rightarrow \infty} \tau_k < \infty$ to prove the global convergence. Here, due to the fact that we only need c_k be bounded, we just require $\tau_k > 0$. A similar strategy has been used in the primal-dual methods of multipliers for complementarity problems [38], where the computational results reported there show that the strategy can make the method quite robust.

6. Computational results

We implement Algorithm 1 in MATLAB and test it on a PC. The examples used here are modifications of the test problem of Marcotte and Dussault [39] and Taji, Fukushima and Ibaraki [40]. The constraint set Ω and the mapping F are taken respectively as

$$\Omega = \left\{ u \in R^5 \mid \sum_{i=1}^5 u_i = 10, u_i \geq 0, i = 1, 2, \dots, 5 \right\}$$

and

$$F(u) = Mu + \rho C(u) + q,$$

where M is a 5×5 asymmetric positive definite matrix and $C_i(u) = \arctan(u_i - 2)$, $i = 1, 2, \dots, 5$. The parameter ρ is used to vary the degree of asymmetry and nonlinearity. The data of this example are given as follows.

$$f(x) = \begin{pmatrix} 0.726 & -0.949 & 0.266 & -1.193 & -0.504 \\ 1.645 & 0.678 & 0.333 & -0.217 & -1.443 \\ -1.016 & -0.225 & 0.769 & 0.934 & 1.007 \\ 1.063 & 0.587 & -1.144 & 0.550 & -0.548 \\ -0.256 & 1.453 & -1.073 & 0.509 & 1.026 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{pmatrix} + \rho \begin{pmatrix} \arctan(u_1 - 2) \\ \arctan(u_2 - 2) \\ \arctan(u_3 - 2) \\ \arctan(u_4 - 2) \\ \arctan(u_5 - 2) \end{pmatrix} + \begin{pmatrix} 5.308 \\ 0.008 \\ -0.938 \\ 1.024 \\ -1.312 \end{pmatrix}.$$

Thus, in our formulation, $u = x$, $y = 0$, $g = 0$, and

$$A = (1, 1, 1, 1, 1), \quad B = 0, \quad b = 10.$$

The parameters used in the algorithm are set as $\mu = 0.6$, $\sigma = 0.001$, $\underline{c} = 0.1$, $\bar{c} = 5$, $t = 0.01$. The stop criterion is set to be

$$\max \{ \|E(w, c_k)\|, \|E(w, c_k)\|/c_k \} \leq \epsilon$$

with $\epsilon = 10^{-6}$. For comparison, we also code the globally convergent Newton method (GCNM) of Taji, Fukushima and Ibaraki [40]. We use the quadratic-program solver quadprog.m from the MATLAB optimization toolbox to perform the projection. We rewrite the subproblem in [40] as a linear complementarity problem (LCP) and solve it by Lemke's complementarity pivoting method [41], which finds a solution of LCP in a finite number of steps. The parameters used in their algorithm are set to be the same as those in [40].

Tables 6.1 and 6.2 report the computational results for $\rho = 10$ and 20, respectively. For simplicity, we denote the proposed method by HEPDM in Tables 6.1 and 6.2.

Table 6.1
Numerical results for $\rho = 10$

Starting pint	Algorithm	Num. of iter.	CPU (s)
(25, 0, 0, 0, 0)	GCNM	7	0.03
	HEPDM	32	0.02
(10, 0, 10, 0, 10)	GCNM	6	0.04
	HEPDM	33	0.02
(10, 0, 0, 0, 0)	GCNM	8	0.03
	HEPDM	26	0.01
(0, 2.5, 2.5, 2.5, 2.5)	GCNM	5	0.03
	HEPDM	21	0.01

Table 6.2
Numerical results for $\rho = 20$

Starting pint	Algorithm	Num. of iter.	CPU (s)
(25, 0, 0, 0, 0)	GCNM	8	0.04
	HEPDM	29	0.02
(10, 0, 10, 0, 10)	GCNM	6	0.04
	HEPDM	34	0.02
(10, 0, 0, 0, 0)	GCNM	8	0.05
	HEPDM	41	0.02
(0, 2.5, 2.5, 2.5, 2.5)	GCNM	6	0.03
	HEPDM	22	0.01

The results in Tables 6.1 and 6.2 indicate that the modified alternating direction method is quite efficient. Though the iterative number is larger than Newton-type method [40], the total CPU time is smaller. Especially, the computational cost at each iteration is much smaller.

7. Conclusions

In this paper, we propose a decomposition method for solving variational inequality problems with separable structure (2) and (3). Under mild assumptions, we show the global convergence of the method, especially, the Fejér monotone to the solution set. Many applications of the decomposition methods showed that a proper sequence of parameters $\{c_k\}$ is crucial to the efficiency. In our algorithm, we just require that the sequence be bounded below from zero and above from infinity, which makes us able to choose the parameters freely and allows us to use improved adjustment strategies with constant parameter, which is important from the viewpoint of computation. To make the method more practical, we solve the subproblems approximately. The accuracy criterion we adopt here is the one proposed by the author [28] for nonlinear proximal point algorithms, which is a modification of the one proposed by Solodov and Svaiter [27] for classical proximal point algorithms.

Acknowledgements

The author would like to thank the anonymous referees for their constructive suggestions, which led to the improvement of the paper.

References

- [1] M. Fukushima, Application of the alternating direction method of multipliers to separable convex programming problems, *Computational Optimization and Applications* 2 (1992) 93–111.
- [2] A. Nagurney, *Network Economics, A Variational Inequality Approach*, Kluwer Academics Publishers, Dordrecht, Boston, London, 1993.

- [3] J. Eckstein, M. Fukushima, Some reformulation and applications of the alternating direction method of multipliers, in: W.W. Hager, et al. (Eds.), *Large Scale Optimization: State of the Art*, Kluwer Academic Publishers, 1994, pp. 115–134.
- [4] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.
- [5] R. Glowinski, P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, in: *SIAM Studies in Applied Mathematics*, Philadelphia, PA, 1989.
- [6] G. Chen, M. Teboulle, A proximal-based decomposition method for convex minimization problems, *Mathematical Programming* 64 (1994) 81–101.
- [7] P. Tseng, Alternating projection-proximal methods for convex programming and variational inequalities, *SIAM Journal on Optimization* 7 (1997) 951–965.
- [8] B. Martinet, Regularization d'inéquations variationnelles par approximations successives, *Revue Francaise d'Automatique et Informatique Recherche Opérationnelle* 4 (1970) 154–159.
- [9] R.T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM Journal on Control and Optimization* 14 (1976) 877–898.
- [10] D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximations, *Computer and Mathematics with Applications* 2 (1976) 17–40.
- [11] D. Gabay, Applications of the method of multipliers to variational inequalities, in: M. Fortin, R. Glowinski (Eds.), *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, North-Holland, Amsterdam, The Netherlands, 1983, pp. 299–331.
- [12] J. Eckstein, Some saddle-function splitting methods for convex programming, *Optimization Methods and Software* 4 (1994) 75–83.
- [13] A. Auslender, M. Teboulle, S. Ben-Tiba, Interior proximal and multiplier methods based on second order homogeneous kernels, *Mathematics of Operation Research* 24 (1999) 645–668.
- [14] A. Auslender, M. Teboulle, S. Ben-Tiba, A logarithmic–quadratic proximal method for variational inequalities, *Computational Optimization and Applications* 12 (1999) 31–40.
- [15] A. Auslender, M. Teboulle, Entropic proximal decomposition methods for convex programs and variational inequalities, *Mathematical Programming* 91 (2001) 33–47.
- [16] R.S. Burachik, A.N. Iusem, A generalized proximal point algorithm for the variational inequality problem in a Hilbert space, *SIAM Journal on Optimization* 8 (1998) 197–216.
- [17] Y. Censor, S.A. Zenios, *Parallel Optimization*, Oxford University Press, New York, NY, 1997.
- [18] J. Eckstein, Approximate iterations in Bregman-function-based proximal algorithms, *Mathematical Programming* 83 (1998) 113–123.
- [19] M.V. Solodov, B.F. Svaiter, An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions, *Mathematics of Operations Research* 25 (2000) 214–230.
- [20] M. Teboulle, Convergence of proximal-like algorithms, *SIAM Journal on optimization* 7 (1997) 1069–1083.
- [21] B.S. He, L.Z. Liao, D.R. Han, H. Yang, A new inexact alternating directions method for monotone variational inequalities, *Mathematical Programming* 92 (2002) 103–118.
- [22] P. Tseng, Applications of a splitting algorithm to decomposition in convex programming and variational inequalities, *SIAM Journal Control and Optimization* 29 (1991) 119–138.
- [23] B.S. He, H. Yang, S.L. Wang, Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities, *Journal of Optimization, Theory and Applications* 106 (2000) 337–356.
- [24] S.L. Wang, H. Yang, B.S. He, Inexact implicit method with variable parameter for mixed monotone variational inequalities, *Journal of Optimization, Theory and Applications* 111 (2001) 431–443.
- [25] B.S. He, H. Yang, Q. Meng, D.R. Han, Modified Goldstein–Levitin–Polyak projection method for asymmetric strongly monotone variational inequalities, *Journal of Optimization, Theory and Applications* 112 (2002) 129–143.
- [26] B.S. He, L.Z. Liao, S.L. Wang, Self-adaptive operator splitting methods for monotone variational inequalities, *Numerische Mathematik* 94 (2003) 715–737.
- [27] M.V. Solodov, B.F. Svaiter, A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator, *Set Valued Analysis* 7 (1999) 323–345.
- [28] D.R. Han, A new hybrid generalized proximal point algorithm for variational inequality problems, *Journal of Global Optimization* 26 (2003) 125–140.
- [29] B.S. He, Inexact implicit methods for monotone general variational inequalities, *Mathematical Programming* 86 (1999) 199–217.
- [30] D.R. Han, B.S. He, A new accuracy criterion for approximate proximal point methods, *Journal of Mathematical Analysis and Applications* 263 (2001) 243–254.
- [31] R.S. Burachik, A.N. Iusem, B.F. Svaiter, Enlargement of monotone operators with applications to variational inequalities, *Set-Valued Analysis* 5 (1997) 159–180.
- [32] K.C. Kiwiel, Proximal minimization methods with generalized Bregman functions, *SIAM Journal on Control and Optimization* 35 (1997) 1142–1168.
- [33] R.S. Burachik, B.F. Svaiter, A relative error tolerance for a family of generalized proximal point methods, *Mathematics of Operations Research* 26 (2001) 816–831.
- [34] J.E. Dennis, R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
- [35] O. Güler, On the convergence of the proximal point algorithm for convex minimization, *SIAM Journal on Control and Optimization* 29 (1991) 403–419.
- [36] B.C. Eaves, On the basic theorem of complementarity, *Mathematical Programming* 1 (1971) 68–75.
- [37] D.P. Bertsekas, On the Goldstein–Levitin–Polyak gradient projection method, *IEEE Transactions on Automatic Control* 21 (1976) 174–184.

- [38] J. Eckstein, M.C. Ferris, Smooth methods of multipliers for complementarity problems, *Mathematical Programming* 86 (1999) 65–90.
- [39] P. Marcotte, J.P. Dussault, A note on a globally convergent Newton method for solving variational inequalities, *Operations Research Letters* 6 (1987) 35–42.
- [40] K. Taji, M. Fukushima, T. Ibaraki, A globally convergent Newton method for solving strongly monotone variational inequalities, *Mathematical programming* 58 (1993) 369–383.
- [41] B.C. Eaves, Computing stationary points, *Mathematical Programming Study* 7 (1978) 1–14.